# Annual Reporting Form for SCEDDBO Projects and Cores

## Center Overview

Period covered by the report: 5/1/2012 – 4/30/2013
EPA Agreement Number: RD83329301-0
Investigators: Marie Lynn Miranda, Allison Ashley Koch, Richard Auten, W. Michael Foster, Alan Gelfand, Pamela Maxson, Evan Myers, Jerome Reiter, Geeta Swamy, Redford Williams
Project Period: Year 6

---

**Objectives of the Southern Center on Environmentally-Driven Disparities in Birth Outcomes (SCEDDBO)**

The central mission of the Southern Center on Environmentally-Driven Disparities in Birth Outcomes is to determine how environmental, social, and host factors jointly contribute to health disparities.  Specific aims of the Center are:

1. *To develop and operate an interdisciplinary children's health research center with a focus on understanding how biological, physiological, environmental, and social aspects of vulnerability contribute to health disparities;*

2. *To enhance research in children's health at Duke by promoting research interactions among programs in biomedicine, pediatric and obstetric care, environmental health, and the social sciences and establishing an infrastructure to support and extend interdisciplinary research;*

3. *To develop new methodologies for incorporating innovative statistical analysis into children's environmental health research and policy practice, with a particular emphasis on spatial, genetic and proteomic analysis;*

4. *To serve as a technical and educational resource to the local community, region, the nation, and to international agencies in the area of children's health and health disparities*; *and,*

5. *To translate the results of the Center into direct interventions in clinical care and practice*.

SCEDDBO leverages and promotes active partnerships among the Nicholas School of the Environment, the Duke University Medical Center, Trinity College of Arts and Sciences, as well as the School of Natural Resources and Environment and the Children's Environmental Health Initiative at the University of Michigan and Durham County Public Health (DCPH) and the Lincoln Community Health Center (LCHC).   The Center brings together the expertise of obstetricians, pediatricians, genetic epidemiologists, spatial statisticians, environmental scientists, social epidemiologists, social psychologists, geographers, and community organizations.

**Synthesis across SCEDDBO**.  **Research Project A: Mapping Disparities in Birth Outcomes** provides population-level research on health disparities in birth outcomes.  Spatially-linking 1.7 million birth records with environmental, social, and host factor data layers allows for population-level analysis of potential co-factors identified in both the clinical obstetrics **Research Project B: Healthy Pregnancy, Healthy Baby: Studying Racial Disparities in Birth Outcomes** and mouse model **Research Project C: Perinatal Environmental Exposure Disparity and Neonatal Respiratory Health** studies. The data from Research Project A is spatially linked in GIS to the data from Research Project B.

The two neighborhood assessments (2008 and 2011) undertaken in Research Project B continue to provide important neighborhood-level environmental and social data to Research Project A.  In addition, the environmental data developed for Research Project A works synergistically with the mouse model work in Research Project C.  For example, the air quality data from Research Project A is being used to further refine experimental dose design in Research Project C. In turn, results from Research Project C regarding experimental effects of multiple environmental agents on fetal growth restriction and postnatal somatic and lung development help point to locations in North Carolina where we are looking more closely at air quality impacts on birth outcomes in Research Project A.

Thus Research Project A is an epidemiological study, while Research Project B is a complementary clinical obstetrics project.  Both projects focus on how combined environmental, social, and host factors shape disparities in birth outcomes.  Research Project B also allows for additional host factor analysis.  Research Project C uses a mouse model system to explore how disparities in exposure and response to exposure initiate and/or enhance disparities in birth outcomes and subsequent neonatal respiratory health.  Like Research Projects A and B, Project C explores the effects of *combined* environmental exposures to prototypical air pollutants common in North Carolina (particulate matter and ozone) and non-chemical stressors on fetal growth restriction, neonatal somatic growth, and subsequent lung development and function.

The synergy among the research projects is facilitated by the GIS and Statistical Analysis (GISSA) Core. The GISSA Core allows for data analysis of the very large amount of data through the use of high-end GIS applications in combination with Bayesian spatial hierarchical modeling and other advanced spatial statistical approaches, thus permitting multi-level analysis. Research Projects A and B both apply a Bayesian spatial hierarchical modeling approach to capture uncertainties in pregnancy outcomes and to elucidate the contributions of economic, sociocultural, and environmental stressors on health disparities in pregnancy outcomes.  State-of-the-art GIS methods allow for sophisticated spatial statistical analyses at highly resolved spatial scales.

The GISSA Core also provides the analysis of the biological response and genetic data generated in Research Projects B and C.  The rich source of social, environmental, and host data in Project B, coupled with sophisticated statistical genetic approaches for identifying gene-gene and gene-environment interactions, provides the opportunity to make important discoveries of how these higher order interactions may be working together to promote or prevent adverse birth outcomes.  By serving as a central clearinghouse for statistical analysis, the GISSA Core tracks outcomes in each project and uses these discoveries to guide the analysis in each of the other projects.

The Community Outreach and Translation Core (COTC) facilitates the communication of findings from all three projects. The COTC continues to communicate the results of the neighborhood assessment to community partners and stakeholders.  In addition, the COTC draws on the GISSA Core to develop materials that communicate the results of the research projects in formats and applications that are immediately accessible to the lay public.

## Research Project A: Mapping Disparities in Birth Outcomes

**Period covered by the report:** 5/1/2012 – 4/30/2013
**EPA Agreement Number:** RD83329301-0
**Investigators:** Marie Lynn Miranda (PI), Alan Gelfand, Pamela Maxson, Evan Myers
**Project Period:** Year 6

---

### Objectives of Research

Project A utilizes the conceptual framework of the "weathering hypothesis," which posits that chronic and persistent stressors lead to accelerated biological aging of women, which in turn accounts for adverse birth outcomes among certain subpopulations. The central objective is to determine whether and to what extent joint exposures to socioeconomic and environmental stressors contribute to racial and ethnic health disparities in fetal growth restriction.

Using a geographically-based nested study design moving from analysis of births for the entire State of North Carolina to six demographically and geographically distinct counties to a single health center and state-of-the-art Geographic Information Systems applications with Bayesian spatial hierarchical modeling and other advanced spatial statistical approaches, the specific aims are to:

1. Spatially link detailed birth record, fetal death certificates, socioeconomic, environmental, tax assessor, community-based, and clinical obstetric data at highly resolved scales for the State of North Carolina from 1990-2003;

2. Refine the concept of fetal growth restriction by a) developing a joint distribution for birthweight and gestation using bivariate modeling for live births and fetal deaths – both separately and jointly, and b) defining it in terms of fetal and infant mortality, rather than percentile cut points; and

3. Determine whether and to what extent differential exposures to both environmental and social stressors help explain health disparities in fetal growth restriction among a) African-American women compared to Non-Hispanic white and Hispanic women, b) Older African-American women compared to younger African-American women, c) Hispanic women compared to Non-Hispanic white and African-American women, and d) Foreign born Hispanic women compared to US born Hispanic women.

This project evaluates a large number of factors in diverse populations, providing broad relevance for birth outcomes across time, space, and demography. Identifying social and environmental factors contributing to fetal growth restriction will improve our understanding of disease etiology and explain the racial disparity in disease incidence, leading to effective interventions against poor outcomes in all population groups.

### Progress Report/Summary of Accomplishments Progress Report/Summary of Accomplishments

Over the past year, the Project A research team has moved forward within small groups to discuss research ideas, review progress of current analysis and identify next steps, and work on manuscript preparation.

A continuing goal is the linking of the detailed birth record data to USEPA $PM_{10}$, $PM_{2.5}$, and ozone monitoring data in order to study the impact of *maternal exposure to air pollution* on birth

weight. To this end, in year 6 we were invited to write a review of air pollution effects on birth outcomes which will be completed in year 7.  We are especially focused on incorporating refined exposure metrics to most effectively characterize meaningful exposures. Significant progress has been made on the relationships between air pollution exposures, socioeconomic status, and birth outcomes. We have extended our methodological work with *spatial downscalers* to conduct an applied analysis on racial and socioeconomic disparities in exposure to air pollution across the State of North Carolina (See Gray et al.). While previous studies of the environmental justice dimensions of air pollution limit analysis of populations living near air quality monitoring stations, we use space-time downscaling methods that we previously developed to output predictive surfaces of ozone ($O_3$) and particulate matter < 2.5 μm in aerodynamic diameter ($PM_{2.5}$) at the census-tract level covering all of North Carolina. This analysis seeks to provide a better understanding of the environmental justice dimension of air pollution exposure across the entire North Carolina population. Moreover, in an additional forthcoming work (see Gray et al.), we link the downscaled output to the detailed birth record in order to examine the joint effects of socioeconomic status and air pollution on birth outcomes, using the highly resolved estimated pollution exposures. The downscaled output allowed us to estimate the association between air pollution exposure and birth outcomes for times and locations where exposure data was otherwise unavailable.

Having linked the North Carolina statewide detailed birth record and educational record databases, we have begun examining the impact of pregnancy-related events and exposures on neurodevelopmental outcomes in early childhood. Two manuscripts have recently been accepted for publication. The first, forthcoming in JAMA Pediatrics, (See Gregory et al. 2013), investigates whether induction and/or augmentation during labor may be associated with autism diagnosis in children in grades 3-8. In this work, we use logistic regression modeling for rare events data to first establish an association between labor induction/augmentation and autism diagnosis and then examine whether the association is robust to controlling for successive sets of potential confounders related to maternal demographics, maternal health conditions, and events of labor and delivery, as recorded in the detailed birth record. The second is forthcoming in Pediatric and Perinatal Epidemiology, examining the joint effect of birth outcomes and maternal prenatal smoking on educational test scores in reading and math (See Anthopolos et al. 2013). This study finds that maternal prenatal smoking may interact with birth outcomes on reading and mathematics test scores, particularly among non-Hispanic white children. Additionally, improvements in birth outcomes, even within the clinically normal range, may be associated with improved academic performance.

Our methodological work on expected performance accruing to *synthesizing categorical datasets*, with the objective of enhancing inference, has been accepted for publication at *Statistical Methodology* (See Berrocal et al., 2012). This work deals with a collection of datasets of varying sizes that are all relevant to a particular scientific question, but which include different subsets of the relevant variables, with some overlap. We synthesize cross-classified categorical datasets drawn from a common population where many of the sets are incomplete (i.e., one or more of the classification variables is unobserved), but at least one is completely observed. The method is expected to reduce uncertainty about the cell probabilities in the associated multi-way contingency table as well as for derived quantities such as relative risks and odds ratios.

**Publications**

Anthopolos, R., Edwards, S.E., and Miranda, M.L. "Effects of Maternal Prenatal Smoking and Birth Outcomes Extending into the Normal Range of Academic Performance in 4[h] Grade in North Carolina, USA." *Paediatric and Perinatal Epidemiology.* Forthcoming.

Berrocal, V., Gelfand, A., Holland, D. 2012. "Space-time Data Fusion under Error in Computer

Model Output: an Application to Air Model Quality." *Biometrics*.

Berrocal, V., Miranda, M.L., Gelfand, A.E., and Bhattacharya, S. "Synthesizing Categorical Datasets to Enhance Inference." *Statistical Methodology.* Forthcoming.

Chang H.H., Reich B.J., and Miranda M.L. "A Spatial Time-to-Event Approach for Estimating Associations between Air Pollution and Preterm Birth." *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 62(2):167-179.

Gray, S., Edwards, S.E., and Miranda, M.L. "Race, Socioeconomic Status, and Air Pollution Exposure in North Carolina." *Environmental Research*. Forthcoming.

Gray, S., Edwards, S., Schultz, B., Miranda, M.L. "Assessing the Impact of Race, Social Factors, and Air Pollution on Birth Outcomes: A Population-based Study*." Environmental Health.* Forthcoming.

Gregory, S.G., Anthopolos, R., Osgood, C., Grotegut, C.A., and Miranda, M.L**.**. "Association of Autism with Induced or Augmented Childbirth in North Carolina Birth Record (1990-1998) and Education Research (1997-2007) Databases." *JAMA Pediatrics.* Forthcoming.

Miranda, M.L., Edwards, S.E., Chang, H.H., and Auten, R.L. "Proximity to Roadways and Pregnancy Outcomes." *Journal of Exposure Science and Environmental Epidemiology*. Forthcoming. PMID: 22805991.

Montagna S, Tokdar ST, Neelon B, and Dunson D. "Bayesian Latent Factor Regression for Functional and Longitudinal Data." *Biometrics*. Forthcoming.

**Research Project B: Healthy Pregnancy, Healthy Baby: Studying Racial Disparities in Birth Outcomes**

Period covered by the report: 5/1/2012 – 4/30/2013
EPA Agreement Number: RD83329301-0
Investigators: Redford Williams (PI), Allison Ashley-Koch, Richard Auten, Pamela Maxson, Marie Lynn Miranda, Jerome Reiter, Geeta K. Swamy
Project Period: Year 6

## Objectives of Research

The central objective of the Healthy Pregnancy, Healthy Baby Study is to determine how the interaction of environmental, social, and host factors contributes to disparities in birth outcomes between African-American and white women in the American South. There are four specific aims:

1. Conduct a cohort study of pregnant women in Durham, NC designed to correlate birth weight, gestation, and birth weight x gestation with environmental, social, and host factors;

2. Develop community-level measures of environmental and social factors by inventorying neighborhood quality and the built environment in partnership with local community groups;

3. Create a comprehensive data architecture, spatially resolved at the tax parcel level, of environmental, social, and host factors affecting pregnant women by linking data from the cohort study and neighborhood assessments with additional environmental and socioeconomic data; and

4. Determine whether and to what extent differential exposures explain health disparities in birth outcomes by applying innovative spatial and genetic statistical methods to:

    a. Identify environmental, social, and host factors that cluster to predict birth outcomes in the entire sample,

    b. Determine whether these clusters are more or less present in African-American versus white populations and quantify the proportion of health disparities explained by differences in cluster frequency, and

    c. Identify environmental, social, and host factors that cluster to predict birth outcomes within the African-American and white sub-samples and compare these clusters across racial groups.

## Progress Report/Summary of Accomplishments

Participant recruitment has been completed, with 1889 women successfully completing participation. Women were recruited from Duke University Medical Center (DUMC) and Lincoln Community Health Center. Demographic data indicate that we were highly successful recruiting women most at risk for adverse pregnancy outcomes, particularly low-income, low educational attainment, and non-Hispanic black women.

The following information was collected from participants in the Healthy Pregnancy, Healthy Baby Study:

- Psychosocial measures include: CES-D, perceived stress, self-efficacy, interpersonal support, paternal support, perceived racism, perceived community standing, pregnancy intention, John Henryism Active Coping Scale, NEO Five Factor Inventory of personality.

- Environmental exposure survey measures include: short survey on fish consumption, smoking pattern and exposure to second-hand smoke, and drinking water source.

- Maternal and neonatal medical record abstraction includes: detailed pre-pregnancy medical and social history, antepartum complications, birth outcomes, and neonatal complications.

- Blood samples for genetic and environmental analysis to assess candidate genes related to environmental contaminant (nicotine, cotinine, cadmium, lead, mercury, arsenic, and manganese) metabolism, inflammation, vascular dysfunction, and stress response.

- Cord blood and placental samples are currently being stored for future genetic analysis and evaluation of activity at the maternal-fetal interface.

We were highly successful in collection of participant-level data as well as biological samples, with greater than 90% attainment of maternal blood sample for genetic and environmental analyses. Collection of cord blood and placental samples, which began in June 2007, has also been successful with approximately 944 delivery samples collected.

All maternal data are georeferenced (i.e., linked to the physical address of the mother) using Geographic Information System (GIS) software. The Healthy Pregnancy/Healthy Baby Study also includes two in-depth neighborhood assessments designed to capture both built environment and community-level social stressors and community resources. In order to increase the participant capture rate (first assessment 40%), we expanded our second assessment area, successfully capturing approximately 70% of the participants. The cohort study and neighborhood assessment data are spatially linked to extensive environmental and demographic data at a highly resolved spatial scale.

The bulk of our focus during the no-cost extension has been on data analysis and manuscript preparation. To that end, we detail our most recent work.

*Psychosocial Indicators.* Analyses have been completed on psychosocial influences on birth outcomes. In order to reduce the number of psychosocial variables, cluster analysis has been performed, resulting in three distinct clusters of women. Cluster analysis on personality was also performed, and a paper is in preparation. A paper examining the relationship between the built environment as measured through the Community Assessment Project and women's psychosocial health was published in year 6 (Messer et al. 2012). Analysis on cadmium exposure and psychosocial stress has been conducted and is being prepared for publication in the coming year. Future analyses will continue with a focus on the concomitant exposure to risk from chemical and non-chemical stressors and resulting pregnancy outcomes.

*Maternal Medical Complications.* Fetal health is not only individually determined, but is also influenced by maternal health and well-being. We continue our emphasis on maternal outcomes. In particular, we have focused on hypertensive disorders during pregnancy. As a first step, we are trying to identify factors that affect maternal blood pressure during pregnancy. In order to make use of the entirety of blood pressure readings collected across the pregnancy, we have considered a variety of statistical approaches, including latent trajectory and sparse functional data models. Our goal is to use environmental, social, and genetic data (such as GRK5 polymorphisms) to predict these blood pressure trajectories. Ultimately we hope these predicted trajectories will aid us in predicting birth outcomes; for example, women with monotone-increasing blood pressure trajectories may exhibit poorer birth outcomes than women with U-shaped curves.

*Statistical Methods Development.*  We have been developing new statistical methodologies designed to improve analysis of the Project B data, as well as to advance statistical analysis more broadly. A paper detailing statistical methodology developed in year 5 for accounting for mid-study changes in measurement scales won the Youden award for the best paper in interlaboratory testing methods this past year (Burgette & Reiter. 2012). These methods were needed because the Project B investigators switched laboratories for measuring blood levels of heavy metals midway through data collection in order to take advantage of finer measurement scales.  Exploratory analysis indicated that the distributions of levels for several exposures were markedly different across the labs, so that analyses based on a simple concatenation of the two labs' data would be biased.  Using the second lab scale as the standard, so that effectively measurements before the lab switch are treated as missing, we developed general purpose methodology for imputing plausible values of the missing exposure measurements.  The methods are based on assumptions about the relative ranks of measurements in the two scales, e.g., a measurement in the 10th percentile in one scale should be at the 10th percentile in the other scale. We implemented this methodology on the Project B data to provide the investigative team with improved data product.

In addition, we developed and implemented methods for finding important predictors in quantile regression when there are a very large number of covariates.  These methods adapted the lasso and elastic net penalties for quantile regression.  We applied the methods on a mid-study sample of women to uncover a previously unreported interaction: women who smoke and who have high blood lead levels tend to have babies with lower birth weights.  An article on this research has been accepted for publication by *Epidemiology* (Burgette et al. 2012).

We developed and implemented methods for using factor analysis models in the context of quantile regression.  The investigative team believes that many of the predictors can be grouped into underlying factors. For example, the Project B data contain several variables that measure maternal stress, and arguably we should connect birth outcomes to the underlying factor of stress rather than its individual indicators.  As another example, the data contain several imperfect indicators of smoking status, and we would like to connect birth outcomes to the underyling factor of true smoking status.  We implemented the model on a mid-study sample of women from Project B, and we found that the smoking factor was a strong predictor of low birth weight.  An article on this research was accepted for publication in *Biometrics* (Burgette & Reiter, 2012).

We also developed statistical methods for the genetic data.  The first statistical innovation involving the genetic data is the adverse sub-population regression (ASPR) for multi-variate outcomes with high dimensional predictors.  The ASPR is a two component latent class model, with the dominant component corresponding to (presumed) healthy individuals and the risk of falling in the minority component characterized via a logistic regression. The logistic regression model is designed to accommodate high-dimensional predictors, as occur in studies with a large number of gene by environment interactions, through use of a flexible nonparametric multiple shrinkage approach. The Gibbs sampler is developed for posterior computation. The method was evaluated with the Project B data and has been published in *Statistics in Medicine* (Zhu et al. 2012).

**Publications**

Chang HH, Reich BJ, and Miranda ML. Spatial Time-to-Event Analysis of Air Pollution and Preterm Birth. *Journal of the Royal Statistical Society Series C.* Forthcoming.

Messer LC, Miranda ML, Maxson P. 2012. The Built Environment and Women's Psychosocial Health. *Journal of Urban Health*. 1-15. PMID:  22907713.

Miranda, M.L., Edwards S., Chang, H., Auten, R. Proximity to Roadways and Pregnancy Outcomes. *Journal of Exposure Science and Environmental Epidemiology*. 23: 32-38. PMID: 22805991.

Zhu, B., Ashley-Koch, A.E., and Dunson, D.B. Generalized Admixture Mapping for Complex Traits. G3 (Bethesda). 2013 Jul 8;3(7):1165-75. doi: 10.1534/g3.113.006478. PMID: 23665878

Zhu B, Dunson D, Ashley-Koch AE. 2012. "Adverse Sub-population Regression for Multivariate Outcomes with High-dimensional Predictors." *Stat.Med*.

**Supplemental Keywords**

Pregnancy, preterm birth, low birth weight, preeclampsia, gestational hypertension, racial disparity, African American, environmental stressors, gene-environment interactions, psychosocial stressors, genes, single nucleotide polymorphisms, genetic admixture

## Research Project C: Perinatal Environmental Exposure Disparity and Neonatal Respiratory Health

**Period covered by the report:** 5/1/2012 – 4/30/2013
**EPA Agreement Number:** RD83329301-0
**Investigators:** P.I.: Richard L. Auten, Co-Inv: W. Michael Foster
**Project Period:** Year 6

---

### Objectives of Research: Specific Aims

1. To determine whether maternal exposure to airborne particulates (PM) and/or ozone (1st hit) restricts fetal growth and/or postnatal growth, and impairs lung development/function in newborn mice;

2. To determine whether PM and/or ozone exposure 're-programs' maternal inflammatory responses;

3. To determine whether postnatal (2nd hit) ozone exposure further impairs postnatal somatic and lung development/function following maternal PM and/or ozone exposures;

4. To determine whether genetic or developmental susceptibility to airway hyperreactivity exacerbates maternal and/or postnatal exposure effects on postnatal somatic and lung development/function.

### Progress Report/Summary of Accomplishments

The effects of prenatal diesel on neurocognitive development appear to be robust: postnatal high fat diet or pre-natal maternal social stress demonstrate synergistic effects on impairment of affective and cognitive behaviors in mice, and the effects on microglial activation appear to be TLR4 dependent. We plan to repeat the studies on the effects on respiratory function to determine if the perinatal pollution effects on lung structure and function in offspring are TLR4 dependent.

### Publications

Block ML, Elder A, Auten RL, Bilbo SD, Chen H, Chen JC, Cory-Slechta DA, Costa D, Diaz-Sanchez D, Dorman DC, Gold DR, Gray K, Jeng HA, Kaufman JD, Kleinman MT, Kirschner A, Lawler C, Miller DS, Nadadur SS, Ritz B, Semmens EO, Tonelli LH, Veronesi B, Wright RO, Wright RJ. The outdoor air pollution and brain health workshop. Neurotoxicology. 2012. 3395)972-84.

Bolton JL, Huff NC, Smith SH, Mason SN, Foster WM, Auten RL, Bilbo SD. Maternal Stress and Effects of Prenatal Air Pollution on Offspring Mental Health Outcomes in Mice. 2013. *Environmental Health Perspectives*. In press.

### Supplemental Keywords

Neuroinflammation, air pollution, maternal stress